CODIB-D-56
12 May 1960

UNITED STATES INTELLIGENCE BOARD

COMMITTEE ON DOCUMENTATION

MEMORANDUM FOR: Committee on Documentation

SUBJECT : Machine Translation

Pursuant to the Twenty-second Meeting of the Committee on Documentation, attached for your information is a copy of my statement to be made on mechanical translation before the House Committee on Science and Astronautics.

STAT

Paul A. Borel
Chairman

MACHINE TRANSLATION

Statement of Paul A. Borel, Assistant Director for Central Reference, Central Intelligence Agency, and Chairman of Documentation Committee of the United States Intelligence Board, before the House Committee on Science and Astronautics, May 13, 1960.

Mr. Chairman and Members of the Committee, the Central Intelligence Agency is pleased to respond to your invitation to outline its views on machine translation.

Our interest in MT dates from 1951, when some of our scientists discussed the possibility of developing an automatic indexing and translating machine with Dr. James Perry, then with the M.I.T. Center for International Studies and now Director of the Center for Documentation and Communications Research, Western Reserve University. After some preliminary work, Dr. Perry and CIA representatives in June 1952 attended a meeting at M.I.T. of linguists, logicians and mathematicians on the subject of machine translation. The principal result of that meeting, which was promoted by Dr. Bar Hillel and supported by the Rockefeller Foundation, was the further stimulation of interest and the realization of possibility in the minds of some of the linguists present.

In the next two years or so, CIA reviewed various proposals, including proposals from M.I.T., the Battelle Memorial Institute, and Georgetown University. Some of these were considered jointly with elements of the Department of Defense.

Our position during that period was that the development of a machine translation capability was highly desirable, and hence that we should

support an MT program. We recognized, however, that such a program had implications which transcended the interests of CIA and those of the intelligence community. We therefore considered it preferable that an organization with broader responsibilities than our own be prevailed upon to take the initiative to push a comprehensive MT program. We identified our immediate need as a usable product, i.e., one which might well be far short of a perfect translation but nevertheless highly useful. In return for an early MT capability to produce a usable product, we were willing to leave the achievement of superior results to a longer range program.

This pragmatic approach was our aim and purpose in 1954. It remains our aim and purpose today.

In early 1955, CIA approached the National Science Foundation, and concurrently ascertained the degree of interest in the Department of Defense. These overtures were directly related to one of a succession of proposals by Professor Leon E. Dostert of Georgetown University. Defense representatives were "all in substantial agreement that, while the Department of Defense does not find it possible to authorize any funds for this project, we will be very much interested in any such device once its feasibility has been firmly established". Negotiations with the National Science Foundation culminated, in early 1956, in an exchange of correspondence between Dr. Alan T. Waterman, Director, National Science Foundation, and Mr. Allen W. Dulles, Director of Central Intelligence. The National Science Foundation agreed "to administer any part of a program of research in machine translation which is agreed by all concerned to be desirable". CIA recognized the need for careful planning and coordination "to insure maximum

progress toward our immediate goal of a machine capability to translate the Russian technical literature".

Two short excerpts from Mr. Dulles' letter of February 29, 1956, are useful to point out some of the broad considerations:

> "I should like to reaffirm the deep interest which we in the intelligence field have in the possibility of translation of Russian language materials, particularly in scientific fields, into English by machine. In addition, many of us feel that the degree of human understanding that could be accomplished if language barriers could be lowered without sacrificing linguistic integrity might well be a major step toward peace."
>
> * * * * * *
>
> "It is our opinion that much is to be gained by the early development of a machine capability for translation. The National Security can be well served if we have available the scientific and technical literature of the USSR in English for detailed analysis as early after publication as possible. I am assured by leaders in electronic research that technological problems yet unsolved need not stand in the way of the rapid development of a machine once the linguistic research had been started."

In the period between May 1956 and the present, the U.S. Government has provided financial and logistical support for the Georgetown project total-

25X1 [redacted]

National Science Foundation

Department of Defense
(estimated value of computer time provided by Air Force and Army)

CIA grants (direct or reimbursement to NSF)

CIA assistance in text preparation

[redacted] 25X1

Considering the inherent difficulties of the undertaking, we have been very pleased with the substantial progress made by Georgetown University.

Efforts to date have been experimental, with Russian organic chemistry texts processed totaling about 500,000 words, and those in [redacted] STAT physics some 200,000. It would be extremely valuable to apply the lessons learned to the automatic translation of several million words of texts covering various diciplines of particular interest, such as, organic chemistry, geophysics (astronomy, meteorology and celestial mechanics), physical chemistry, high energy physics and solid state physics. And in this connection we are currently studying a proposal by Georgetown University to conduct a large scale operational feasibility test during the next fiscal year. As a further indication of the development in the state of the art, I should also mention that only last week we received a proposal by the International Business Machines Corporation for the establishment of an automatic translation facility. The Committee may be interested in the exchange of correspondence between Dr. E. R. Piore, Director of Research, IBM, and Mr. Dulles, which I will be pleased to make available. This proposal too is under study.

It is fair to ask whether developments since our initial interest in MT have called for a change in original objective. In the main, the answer is no. In recent years the volume of available Russian scientific and technological literature has greatly increased. We estimate that the

25X1 [redacted] This increase has been accompanied by increased efforts by the Government to translate the most useful part of this production. And the performance, 95% of which is by the Government or under Government contract, is impressive. About 53 million words of Russian scientific literature are now being translated annually (of which CIA accounts for over 9 million).

Why then the need for MT? The reasons can, I think, be simply stated:

(1) The volume of publication will continue to increase, and at a rate in excess of our ability to procure competent translations.

(2) The quality of translation work done through contract arrangements is not uniformly excellent. Whatever the level of accomplishment in MT at any given time, the output is uniform. In short, MT holds out the promise of a uniformly more accurate product.

(3) MT also promises greater speed. We now give priority to categories and languages of greatest interest. Non-priority items are invariably slow in reaching the reader. Perhaps the translator with the particular skill in a language, or in a discipline, cannot immediately take on the task. In any case he cannot translate on an average [redacted] 25X1
The machine can hurtle them out at rates of [redacted] 25X1
[redacted] And 25X1
these rates will increase. Even if post-editing were required, the man-machine system would appreciably out-produce the human translator working alone.

(4) With MT, more translations would be available. This increased availability of translations would itself generate new and more widespread demands for them. We now strive to pass over only marginal material, but cannot be sure that we are invariably successful.

(5) Greater availability would result in a better informed corps of scientists in this country. This would result in superior evaluations of scientific and economic developments in the Bloc than is now possible.

(6) The development of a two-way MT capability would make possible low cost production of American publications for sale in underdeveloped countries where low cost Bloc publications now have an almost clear field to the detriment of U.S. interests.

(7) And finally, the research done and the techniques developed for accomplishing translation by machine would contribute materially to the solution of problems in the broader field of information storage and retrieval, and the emerging field of language data processing.

A word should be said about some of the problem areas. I am not too concerned about the technical problems. There are many of course. For example, an all purpose rapid print reader would certainly be essential to an efficient system. Today we prepare computer input by punching each Russian word onto IBM cards and then go from card to magnetic tape. Our

[redacted] 25X1

manner. But good work is being done in the technical area and solutions will be forthcoming. More basic is the problem of organization. Shall an MT capability once achieved be exploited by each on his own or should a central facility serve all? If the latter, who shall set it up, who shall operate it, and under what terms shall Government and private interests participate?

It is not too early to start thinking about this. A central facility is, I believe, indicated but not exclusively so. The enormous potential output of MT greatly exceeds the present and prospective requirements of any one part of Government or single private organization. Problems of procuring and selecting materials to be translated, and of disseminating translations to those needing them, are very considerable and can most efficiently and economically be solved centrally. Moreover, the use of a central facility permits the use of equipment exclusively designed to produce automatic translations. There are, however, requirements for accomplishing translations under mobile conditions, or, for fully utilizing general purpose equipment acquired for processing data rather than lexical material. Hence there is also continuing need for research to develop multiple language, multiple discipline MT materials and programs for translation by general purpose computers.

When MT is discussed there is invariably an expressed interest in what the Soviet Union is doing in this field. I will not dwell on this except to say that the Soviets have a program which considerably exceeds our own in scope and size, and that they are doing very good theoretical work, though restrictions on the availability of computer time has limited opportunities to apply theory to practice. Two papers, one by Professor Oettinger 1/ and one by Dr. Harper 2/, provide valuable assessments of the Soviet effort. The Joint Publication Research Service series Soviet Developments in Information Processing and Machine Translation, will also be of interest to the Committee.

1/ Anthony G. Oettinger, "A Survey of Soviet Work on Automatic Translation", Mechanical Translation, Vol. 5, No. 3, December 1958, pp. 101-110.

2/ K. E. Harper, "Soviet Research in Machine Translation", RAND Corporation Monograph No. P-1896, 4 Feb. 1960, 17 pp.

I will close my remarks with a word about coordination of MT activities among Government departments. Until recently only a few departments or agencies had programs. As the field grew CIA took steps to formalize within the intelligence community the informal channels of communications used by those in charge of MT programs, whether in the intelligence or the research components of their respective departments. A group of experts was formed whose membership is drawn from the Army, Navy, Air Force, State, NSA, and CIA, with a National Science Foundation representative as associate member. The function of this interdepartmental group is: to advise the Committee on Documentation of the United States Intelligence Board with respect to all machine translation activities, to coordinate MT activities within the intelligence community, and to inform its members of new projects and of the status of existing projects. For over-all coordination in matters transcending the interests of the intelligence community, CIA looks to the National Science Foundation.

C
O
P
Y

INTERNATIONAL BUSINESS MACHINES CORPORATION

590 Madison Avenue,
New York 22, N.Y.

Office of the Director of Research

April 20, 1960

Mr. Allen W. Dulles
Director,
Central Intelligence Agency
Washington 25, D. C.

Dear Mr. Dulles:

Last February IBM had the opportunity to present its research program on automatic translation of languages to Mr. Amory and other members of the Agency's staff. Following the presentation a number of suggestions were made by members of your staff that we should consider seriously the establishment of an Automatic Translation Facility.

Apart from the research and development program, we have made studies of the operational requirements, and following the suggestions we have coordinated them into a proposal, which we would be very pleased to present to the Agency at your convenience.

A large part of our program is supported by the USAF Air Research and Development Command. We have received official permission from ARDC to make this presentation, should you wish to consider it at this time, and to discuss the possibility of consolidating their program with the Agency's, in the interest of initiating a national Automatic Translation Facility.

S/ E. R. Piore

E. R. Piore
Director of Research

ERP:jms

C
O
P
Y

CENTRAL INTELLIGENCE AGENCY
Washington 25, D.C.

OFFICE OF THE DIRECTOR

30 April 1960

Dr. E. R. Piore
Director of Research
International Business Machines Corporation
590 Madison Avenue
New York 22, N.Y.

Dear Dr. Piore:

Thank you for your letter of 20 April, in which you offer to present a proposal for the establishment of an Automatic Translation Facility.

We have since 1952 urged the development of a machine translation capability, particularly for translating scientific Russian to English. Since 1956, we have provided financial support for the work done in this field at the School of Languages and Linguistics of Georgetown University.

We believe that the progress made to date by Georgetown University and others working on various aspects of this difficult problem is sufficiently encouraging to justify a large scale operational feasibility test.

It is therefore most timely to explore prospects for consolidating programs, and we are pleased to learn of Air Force interest in initiating a national Automatic Translation Facility.

Since it would be helpful to us to review your proposal at an early date, Mr. Amory has been in touch with Mr. John Griffith and a meeting has now been arranged for May 5.

Sincerely,

S/ Allen W. Dulles

Allen W. Dulles
Director

CODIB-D-55
12 May 1960

U N I T E D S T A T E S I N T E L L I G E N C E B O A R D

COMMITTEE ON DOCUMENTATION

MEMORANDUM FOR: Committee on Documentation

SUBJECT: Sub-Committee on Classification

1. Pursuant to paragraph 7 of the Minutes of the Twenty-first Meeting, 31 March 1960 (CODIB-M-21, dated 6 April 1960), Mr. Louis C. Covell, Chief, Document Division, CIA is designated Chairman of the Classification Sub-Committee and Mrs. Josephine R. Brahm, Staff Assistant to Chief, Document Division, CIA is designated Vice-Chairman.

2. The Classification Sub-Committee will be concerned primarily with establishing procedures for revising the Intelligence Subject Code and for approving any changes, deletions and additions thereto. In addition, the Sub-Committee will be asked by CODIB to study problems related to the indexing of intelligence documents and submit recommendations for the establishment of uniform procedures.

3. The Sub-Committee Chairman has asked that the names of the persons who will represent the USIB Agencies on the Sub-Committee be submitted so that an organizational meeting can be called at an early date.

Paul A. Borel
Chairman